

System and Method for Caching and Rendering Images**Cross Reference to Related Applications**

5 This application claims priority to and benefit of United States Provisional Patent Application 60/394,412 entitled "System and Method for Caching and Rendering Images" filed July 5, 2002 by Andrew Cohen and Scott Crouch, the entire contents of which are hereby specifically incorporated by reference for all it discloses and teaches.

10 **Background of the Invention**

a. Field of the Invention

The present invention pertains generally to rendering digital images and specifically to high speed rendering vector and raster images using caching.

15 b. Description of the Background

Maps and other large graphical databases are difficult to display and manipulate on smaller computing devices. For example, hand held computers have small displays but also slow central processing units and limited memory. For these devices to be useful to display and manipulate image databases, the speed and memory limitations must be overcome. Other computing devices have similar memory and computational power limitations and can also benefit from enhanced performance.

20 Map data comes in two general forms: raster and vector data. Raster data may be bitmap images, such as aerial photographs and topological maps. Some map data may be in the form of vector data. For example, political boundaries or the routes of the interstate highway system may be in vector form. The generation of images may 25 comprise both raster and vector data.

25 Databases that contain large images are often terabytes in size whereas the computing devices, such as hand held computers, may have only megabytes of storage space. Only a small portion of the database may be stored locally and used at any one time. In many cases, when a user pans or zooms the image, they may have to wait for a

new portion of the database to be loaded and displayed on the screen. Sometimes, the 'lag time' for the new information to load is noticeable and frustrating for the user.

It would therefore be advantageous to provide a system and method for caching and rendering image databases that is fast and uses memory efficiently. It would further 5 be advantageous if the system and method used the displaying computer's computational power efficiently.

Summary of the Invention

The present invention overcomes the disadvantages and limitations of the prior art 10 by providing a system and method of preprocessing a raw image, then displaying the image using a caching system wherein the caching system predicts and loads tiles of the preprocessed raw image that may be useful to the user. The system may be adapted to both raster and vector data.

The present invention may therefore comprise a method of caching and rendering 15 an image comprising: providing a raw image to render; preprocessing the raw image to produce a plurality of tiles; determining a starting view, the starting view being a portion of the raw image and comprising a first set of tiles; loading the first set of the tiles into memory; displaying the starting view using the first set of tiles; providing a weighting algorithm that determines a weight for tiles; applying the weight to each of the tiles based 20 on the weighting algorithm; providing a first analysis algorithm capable of analyzing the weight of each of the tiles to identify tiles that are not in the memory; identifying at least one tile that is not in the memory using the analysis algorithm; and loading the at least one tile into the memory.

The present invention may further comprise a method of caching and rendering a 25 raw image comprising: receiving a raw image to render, the raw image comprising a plurality of vectors; determining a starting view to display, the starting view being a portion of the raw image; determining a cache space that is a superset of the starting view; determining a first set of intersecting vectors that intersect the cache space; trimming at least one of the vectors in the first set of vectors to approximately the limits 30 of the cache space; storing the intersecting vectors in a cache; rendering the starting view by displaying the first set of intersecting vectors from the cache; receiving a first

command from a user to change the view to a second view; determining if the second view is completely within the cache space; if at least a portion of the second view is within the cache space, rendering the second view using at least a portion of the intersecting vectors stored in the cache; and if at least a portion of the second view is

5 outside of the cache space, determining additional intersecting vectors that intersect the second view and rendering the second view using at least a portion of the additional intersecting vectors.

The present invention may further comprise a system for caching and rendering an image comprising: a first computer system adapted to preprocess the image to produce

10 a plurality of tiles; a second computer system adapted to determine a starting view, the starting view being a portion of the raw image and comprising a first set of tiles, load the first set of the tiles into memory, display the starting view using the first set of tiles, apply a weight to each of the tiles based on a weighting algorithm, identify at least one tile that is not in the memory using an analysis algorithm, and load the at least one tile into the

15 memory.

The present invention may further comprise a system for caching and rendering a raw image comprising: a first computer program adapted to receive a raw image to render, the raw image comprising a plurality of vectors, determine a starting view to display, the starting view being a portion of the raw image, determine a cache space that

20 is a superset of the starting view, determine a first set of intersecting vectors that intersect the cache space, trim at least one of the vectors in the first set of vectors to approximately the limits of the cache space, and store the intersecting vectors in a cache; a second computer program adapted to render the starting view by displaying the first set of intersecting vectors from the cache, receive a first command from a user to change the

25 view to a second view, determine if the second view is completely within the cache space, if at least a portion of the second view is within the cache space, render the second view using at least a portion of the intersecting vectors stored in the cache, and if at least a portion of the second view is outside of the cache space, determine additional intersecting vectors that intersect the second view and rendering the second view using at

30 least a portion of the additional intersecting vectors; a first computer adapted to execute

the first computer program; and a second computer adapted to execute the second computer program.

The advantages of the present invention are that potentially useful portions of an image are identified, cached, and ready to be displayed for a user. The determination of 5 which sections of a large image database to cache is based on rules and algorithms that may be adapted to particular applications. Both raster and vector image databases may be used with the present invention. Using the present invention, the speed of the display is not affected by the amount of data to be displayed.

10

Brief Description of the Drawings

In the drawings,

FIGURE 1 is an illustration of the connection of the Internet with a desktop computer and a hand held computer.

15

FIGURE 2 is an illustration of a workflow diagram of an embodiment of the present invention of a method of rendering a portion of an image.

FIGURE 3 is an illustration of a workflow diagram of an embodiment of the present invention of a method for weighting and caching tiles that are displayed.

20

FIGURE 4 is an illustration of a method for weighting tiles and an example of an embodiment of the decision criteria for adding new tiles and discarding old tiles from the cache.

FIGURE 5 is an illustration of a work flow diagram of a method of an embodiment of the present invention for handing vector data in a cached space.

FIGURE 6 is an illustration of the concept of trimming vector data that extends outside of cache space.

25

Detailed Description of the Invention

Figure 1 illustrates the connection of the Internet 102 with a desktop computer 104 and a hand held computer 106. In one embodiment of the present invention, the desktop computer 104 may receive an image from the Internet 102, preprocess the image, 30 and transfer the image to the hand held computer 106.

The present invention allows large, complex images, such as maps, to be preprocessed and then displayed. The display device may have a slow processor and limited memory that makes displaying and panning and zooming of large databases almost impossible. The present invention allows for the efficient use of memory and 5 relatively slow processing speed on smaller devices to make the image viewing experience of a user similar to a high speed device.

The present invention is directed at the display of large images on computing devices. For example, a map image from a server on the internet 102 may be prepared by 10 a desktop computer 104 and transferred to the hand held computer 106. The hand held computer 106 may be detached from the desktop computer 104 and used in the field as a portable device.

In another embodiment of the present invention, a web server may contain a database of images. The web server may preprocess an image and prepare it to be displayed on a desktop computer using the present invention. In other embodiments, 15 images may be displayed on global positioning system receivers, cellular phones, automotive computers, wireless devices, or the like.

In further embodiments of the present invention, a computer aided design system may display a portion of a mechanical design or a game system may display images associated with a game.

20 Figure 2 illustrates a workflow diagram of an embodiment of the present invention of a method 200 of rendering a portion of an image. The raw image is loaded 202, preprocessed 204, an origin or first view is chosen 206, tiles for the first image is loaded 208, and the user may zoom or pan the image 210.

The preprocessing 204 may comprise converting the image to pixel or screen 25 space 212, converting the image to a format compatible with writing directly to the display 214, optionally subdividing the raw image into tiles 216, and other preprocessing 218.

The raw image may come from virtually any source and contain any type of data. 30 Map images, CAD data, or other graphical image may be used. The image may contain either or both raster or vector data.

The preprocessing 204 may or may not be performed on the same computing device that performs the pan and zoom 210. In some embodiments, the preprocessing 204 may be done at the time a user requests an image. Such a system may have a lag time wherein several functions are necessary to prepare an image to be displayed. In 5 other embodiments, images may undergo the preprocessing 204 and be stored in a database prior to the user's request for the image.

The preprocessing routine 204 prepares an image for display on a device. For the example of a hand held computer, a map image may be converted from map coordinates to pixel or screen space 212. The step 212 may involve converting an image from 10 elliptical geophysical coordinates to the rectangular coordinates of a display.

The image may be converted to a format compatible with writing directly to the display 214. For example, if the image contained 256 colors and the display was capable of displaying 16 colors optimally, the color of the image may be converted. To optimize the speed of the image panning and zooming, it may be desirable to write the image 15 directly to the display hardware and bypass any software processing that a computer operating system may perform. Such techniques are common for graphics intensive applications.

Subdivision of the raw image into tiles 216 may optionally be performed at this time. In order to handle the large image, it is subdivided into tiles. If the initial size of 20 the first view is known, the image may be subdivided during the preprocessing step 204. In other cases where the size of the initial view or subsequent views are not known, the subdivision of the image into tiles may be performed after the first view is determined.

The user may determine the first view of the image 206 through many mechanisms. For example, the origin of an image may be selected through a software 25 application such as a website that is offering directions to a particular destination. In other cases, the user may choose a portion of a larger image to view by indicating the initial view from a scaled down view of the full image. The first view of the image 206 may include the origin or center and the scale or extents of the view. In some cases, the user may select two corners of the view.

30 The tiles are loaded 208 to display the first view of the image. The tiles may be sections of the image that, when butted up against each other, appear to be seamlessly

continuous. The tile size may vary for different applications and different processing capabilities of the display device.

The user may zoom or pan the image 210 to change views during the course of using the image data. In some cases, the user may manually select the portions of the 5 image to be displayed. In other cases, the device may change views under software control.

Figure 3 illustrates a workflow diagram of an embodiment of the present invention of a method 300 for weighting and caching tiles that are displayed. The method 300 incorporates the method 200 illustrated in Figure 2.

10 The raw image is loaded 302, preprocessed 304, a starting view is determined 306, tiles are created from the raw image 308 if tiles were not created during the preprocessing step 304, and the tiles are loaded into the initial view 310. When the tiles are loaded into the initial view, the tiles are also loaded into a cache, whereupon the tiles are assigned weights 312 and the view is displayed 314.

15 The weights assigned to tiles are used by an algorithm for maintaining certain tiles in a cached memory for quick retrieval. Further, the weights of various tiles may be analyzed to predict which tiles may be used in the future and thus may be also loaded into the cache. The cache is an area of memory that is set aside by the present invention for holding tiles to be displayed. The cache may be a general area of memory or may be a 20 dedicated memory specially adapted to cache tiles for very fast display.

When a zoom or pan command is issued 316, the display is updated to the required view 318. The tiles in cache are used and any tiles not in the cache are retrieved and brought into cache to be displayed. The tiles are assigned new weights 320, the analysis thread is launched 322, and the process awaits another pan or zoom command 25 316.

The analysis thread 324 is done during the period of time when the user is viewing the new view displayed on the screen. During this time, the computing device that is displaying the view may have some unused processing cycles whereby the cache may be updated.

30 The analysis thread 324 comprises deleting old, unused tiles 326, analyzing the cache to find new tiles to add 328, adding the new tiles 330, determining if the number of

tiles is optimal 332, and resizing the tiles 334 if necessary. The analysis thread 324 manages the cache to optimize the tiles that are likely to be used. By keeping the correct tiles in the cache, the user may zoom and pan while enjoying very fast response from the display device.

5 The method for weighting the tiles, selecting those tiles to be deleted from cache, and determining which tiles to add to the cache may take on several forms. The characteristics of the weighting methods, including the criteria for selecting new tiles and discarding old tiles, may be dependent on the various applications. For example, the characteristics and behaviors of a user of a map displayed on a GPS receiver may be

10 10 different from an architect or contractor viewing an image of building plans. Thus, the particular algorithms suited for each application may be different. Those skilled in the arts may use various weighting algorithms and criteria for adding and discarding tiles while still maintaining within the scope and intent of the present invention.

15 The size of the tiles may be adjusted 334 depending on several factors. For example, when the user zooms out, many more tiles are displayed on the screen at one time. The increased number of tiles means more computations are required to analyze the cache in addition to other factors. For some applications where the user may zoom in and out of the image repetitively, the tiles may be less likely to be resized at each zoom command. In such cases, the tiles may only be resized if the zoom ratio was relatively

20 20 constant for a period of time.

25 Figure 4 illustrates a method for weighting tiles and an example of an embodiment of the decision criteria for adding new tiles and discarding old tiles from the cache. Four sequential views 402, 404, 406, and 408 are shown. Each view is comprised of many individual cells 403, each of which represents a tile that may be placed in the cache.

30 In the present example, the weighting algorithm is as follows: each tile that is displayed has its weight incremented by one and those not displayed are decremented by one. Those tiles that were displayed in the previous view, but are not displayed in the current view, are not decremented but keep their weight unchanged. Those tiles with a weight of zero are removed from the cache.

In the first view 402, the displayed tiles 410 are all incremented to one by virtue of being displayed for the first time. At this point, only the tiles displayed are loaded into the cache. In other algorithms, additional criteria may be used to load additional tiles into cache. For the purposes of illustration, such criteria are not included in the present 5 example.

In the second view 404, the user has panned the view downward one tile and rightward one tile. All of the tiles in the displayed tiles box 412 are incremented by one. The unused tiles from the previous view are left unchanged and are kept in the cache. Several predicted tiles 416 are loaded into the cache as well.

10 The prediction algorithm may analyze the movement of the view from the previous view. For example, the algorithm may determine the direction of movement and add those tiles necessary to display an additional movement continuing in the same direction.

15 The prediction algorithm may analyze the topology of the weights in the cache and determine tiles to add to the cache based on the slope or change of slope of successive tiles. For example, if successive tiles had weights of 2 and 1, the analysis may predict that the next tile in succession may be added to the cache. Such analyses may be performed horizontally, vertically, and diagonally across the cache.

20 In some cases, the predicted tiles 416 may be given a weight of one. In other cases, the predicted tiles 416 may be loaded and given another weight, such as 0.5, or no weight at all. The particular algorithm and rules used to predict which tiles to add may be changed based on the particular needs and behaviors of the user. Further, the capabilities of the display device may also be a factor when determining the particular algorithm and rules. For example, a device with very limited computational power may be capable of 25 only rudimentary analysis whereas a high powered system may be capable of executing more detailed analysis algorithms.

30 The size of the cache may be a factor in the algorithm and rules for selecting which tiles to add to the cache. For example, a large cache may allow for an algorithm to add two, three, or more tiles past the limits of the displayed tiles 412 in the direction that the user has moved the view.

In the third view 406, the displayed tiles 418 are again incremented and several tiles are discarded 420. Additional tiles are predicted 422 and added to the cache. In the present view 406, all of the displayed tiles 418 were contained in the cache in the second view 404. In the third view 406, the cached tiles would be very quickly loaded and

5 displayed.

In the fourth view 408, the user has moved the view downward one tile and leftward one tile. The displayed tiles 424 are incremented, several tiles are discarded 426, and several tiles are predicted 428 and thus added to the cache. When the user executed the move to the fourth view, all of the tiles were in memory except the missing

10 tile 430. Tile 430 may have needed to be added to the cache to complete the display and may have caused a delay in completing the view requested by the user.

The algorithms and rules for discarding various tiles from the cache may include the effects of the limitations on memory size of the device. In the present example, all tiles that have a weight of zero are eliminated. In some embodiments, those tiles may be

15 kept in the cache until the cache is close to being filled, at which time the tiles with the lowest weight would be removed to make room for new tiles.

Those skilled in the arts may develop various algorithms for selecting tiles to add and delete to and from a cache while maintaining within the scope and intent of the present invention. The example contained herein was selected to best illustrate a simple

20 algorithm whereby a cache may be maintained. Various mathematical tools, including statistical analyses, adaptive algorithms, rules based on historical behaviors, and other algorithms may be developed as necessitated by the various applications of the present invention.

Figure 5 illustrates a work flow diagram of a method 500 of an embodiment of the present invention for handing vector data in a cache space. When the initial view is

25 loaded 502, the extents of the cache space is determined 504, vectors that cross cache space are found 506 and trimmed 508. After a zoom or pan command is issued 510, cache space is adjusted 512 and any adjustments to vectors in cache space are made 514. The process continues by finding new vectors that cross cache space 506 and trimming

30 those vectors 508.

The term cache space is used to denote that portion of the larger image that is in the cache at any time. Cache space is a subset of the raw image, but a superset of the displayed image.

5 The method 500 may be used for purely vector images or in conjunction with the method 300 to handle raster and vector data in the same image. For example, a map image may incorporate raster images with vector data for roads and political boundaries. Both types of data may be used in the present invention.

Determining the extents of cache space 504 and adjusting cache space 512 may be done using the various algorithms and rules as described for the method 300. In some 10 cases, the rules for vector data may be adjusted for the peculiar nature of vector data. For example, a small shift in the cached space limits may not require that trimming or extending of a vector. In such a case, the computations required to further trim a vector may be avoided until a more significant shift in the cache space has occurred. As with method 300, various algorithms and rules may be developed for each application 15 whereby the vector elements may be trimmed and extended based on a shift in cache space.

Figure 6 illustrates the concept of trimming vector data that extends outside of cache space. The raw image space 602 contains the displayed portion of the image 604 and the cached area 606. A polyline 608 extends through the cached space 606 and the 20 displayed portion 604. The endpoints 610 and 612 are the nearest endpoints on polyline 608 that are outside of the cache space 606. During a trimming process, the sections 614 and 616 are discarded from the polyline 608 and are not stored in cache memory.

Similarly, polygon 618 extends outside of the cache space 606. A temporary boundary line 622 is calculated between endpoints 624 and 626 and the section 620 of 25 polygon 618 is discarded from memory.

As the cache space 606 moves due to shifts in the displayed image 604, the limits and extents of the polyline 608 and polygon 618 may also be adjusted using the method 500.

Hence, the present invention improves the image display performance by caching 30 certain portions of an image in a high speed memory so that subsequent views of the image can be updated quickly. The cached portions of the image may be preprocessed

and ready for directly writing to a device's display memory for maximum display speed. The present invention may allow limited capacity devices such as hand held computers and cellular phones to perform panning and zooming of large images with performance that rivals high performance computers. Further, the present invention may further 5 enhance the performance of desktop computers.

The caching strategy has the ability to identify sections of an image that may be viewed in the future. These sections may be identified by algorithms and rules that may be tailored to specific applications and user behaviors. The ability to 'look ahead' and predict a user's needs means that the computational and time intensive operations of 10 finding a portion of an image, reading the portion from a storage device, and preparing the portion to be displayed is done during periods when the device is waiting for user input and is otherwise idle.

The foregoing description of the invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to 15 the precise form disclosed, and other modifications and variations may be possible in light of the above teachings. The embodiment was chosen and described in order to best explain the principles of the invention and its practical application to thereby enable others skilled in the art to best utilize the invention in various embodiments and various modifications as are suited to the particular use contemplated. It is intended that the 20 appended claims be construed to include other alternative embodiments of the invention except insofar as limited by the prior art.